



Tail probabilities for triangular arrays

Citation

Fudenberg, Drew, and David K. Levine. 2013. Tail Probabilities for Triangular Arrays. *Journal of Dynamics and Games* 1(1): 45–56. doi:10.3934/jdg.2014.1.45

Published Version

doi:10.3934/jdg.2014.1.45

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:13041349>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Note: this version still has some typos we cleaned up later so the journal version is recommended.

Tail Probabilities for Triangular Arrays

Drew Fudenberg and David K. Levine¹

Departments of Economics,
Harvard University and Washington University in St. Louis

First Version: November 1, 2007

This version: September 10, 2014

¹ We thank Yuhta Ishii and Mihai Manea for careful proofreading and the associate editor and referees for very helpful comments and suggestions. NSF grants SES-03-14713, SES-04-26199, SES-0646816, SES-0851315, and SES-0954162 provided financial assistance. We are grateful to Tom Ferguson and Tom Liggett for suggestions about the literature and on how the Feller proof works.

1. Introduction

It is frequently difficult to determine the set of equilibrium payoffs in discrete time repeated games with imperfect public monitoring when the discount factor is bounded away from one. In the continuous time case Sannikov [2007] and Sannikov and Skrzypacz [2007] have obtained striking characterizations of the equilibrium set in continuous time games where the public signals are modeled as a diffusion process, with the players' actions altering the diffusion's drift but not its volatility. These continuous-time models are motivated as modeling the limit of very high frequency interactions, which raises the question of what sorts of high-frequency limits the models capture. This in turn depends on the relationship between the signal processes in discrete and continuous time. Fudenberg and Levine [2009] (hereafter referred to as FL) show by example that the same limiting diffusion processes can arise as the limit of different discrete-time structures that have very different limit equilibria.

In characterizing the “cooperative” equilibria of a repeated game it is necessary to understand which “punishment schemes” are incentive compatible for players. This can be thought of as testing for whether a deviation has occurred combined with a punishment if the test is failed. Intuitively, as with the normal distribution, the tails of a diffusion process permit a very accurate test for the difference in means by using a cutoff for the signal, above which the test is considered to have “failed.” However, since the worst possible punishment in a repeated game is bounded, what matters is not just the accuracy of the test but whether defections can be detected with sufficient probability. As we approach continuous time as the limit of shorter discrete intervals, the question becomes how rapidly the probability with which defections can be detected decreases relative to the size of available punishment. If the only way to create a sufficiently accurate test is to send the cutoffs very quickly to infinity, then punishment will occur too rarely to provide sufficient incentives for cooperation. In this case we can expect that there will only be static equilibria in the limit. Consequently a key question is whether it is possible to design a test that finds an appropriate balance between accuracy and frequency of punishment as the period length shrinks. For concreteness we will illustrate this idea in a simple principal-agent game instead of the repeated game studied in FL.

In many – if not most – cases of interest, the public signal is not literally continuously distributed, but the diffusion process arises as the limit of the aggregate of many small discrete events such as price changes. In this case we are interested not in the normal distribution per se, but rather a distribution that approaches normality in the limit. It might be hoped that a version of the central limit theorem could be used to examine the convergence properties of the test statistic. Unfortunately as periods shrink the optimal cutoff increases in such a way that the probability of detection decreases (the cutoff normalized by the standard deviation increases) so the standard central limit theorem is not useful. Instead what is required is an estimate of the tail probabilities, that is of the probabilities of very unlikely but informative signals.²

The most closely related result in the literature is what Feller [1971] calls a “large deviations” theorem, although that term is now used for other things. Feller’s result applies only to i.i.d. random variables, and not to triangular arrays; this note provides the additional uniformity assumptions needed to adapt the Feller proof to the case of triangular arrays and adapts the proof to show how these uniformity assumptions are used. The result reported here can then be used to show that the equilibria of discrete-time games whose signals are binomial arrays do indeed converge to the equilibria of the associated continuous time game, as it was in FL’s study of games with a long run player against a myopic opponent. In the next section we sketch a simpler one-shot agency problem where the tail probability estimates can be used in similar way.³

2. A Motivating Example

The information issues that arise in repeated game setting arise in a simplified form even in a principal-agent problem, as we now show. Suppose that there is a period of length τ . At the beginning of the period the agent may choose not to be employed by the principal in which case he receives zero. If he chooses employment he must decide between *working* (W) and *shirking* (S). If he works he is paid an amount $W\tau$

² This issue is delicate because the likelihood ratio between signals between two normal distributions with a common variance and different mean becomes unbounded in the tail: this was originally exploited by Mirlees [1974].

³ Sadzik and Stachetti [2012] study the limit of discrete-time agency problems when the discrete-time signals have a continuous density as opposed to being the sum of discrete random variables. Their “hidden action” case corresponds to the example presented here.

proportional to the length of time he works. If he shirks he gets a bonus of $G\tau$. At the end of the period, a principal observes a noisy signal y of the agent's lack of effort and if this signal exceeds a threshold \bar{y} he imposes a fixed penalty P . Notice that P is not proportional to the length of the period; the idea is that the principal can impose a long-term punishment on the agent if he feels the agent has shirked even for a short period of time. For example if the principal can fire the agent, then we would expect that $P = W/r$, which is the amount that the agent would have earned from a lifetime of employment with the principal.

The question we wish to address is for particular distributions of y whether it is possible to set the threshold \bar{y} so that the agent can be induced to work rather than shirk. Notice that whether or not it is desirable to do this depends on payoffs to the principal which we do not specify.

Let p represent the probability that the punishment is received if the agent works and q the probability of punishment if the agent shirks. Then if it is to be optimal for the agent to work rather than shirk then it should be that the incentive constraint

$$\rho(\tau) \equiv \frac{q - p}{\tau} \geq \frac{G}{P}.$$

holds. This is similar to (1) in FL. If it is to be optimal to choose employment then the participation constraint should be satisfied, that is

$$\mu(\tau) \equiv \frac{p}{\tau} \leq \frac{W}{P}.$$

If in the limit as $\tau \rightarrow 0$ both of these are to hold for some values of G, P, W then it must be that $\lim \rho(\tau) > 0$ and $\lim \mu(\tau) < \infty$. This is analogous to Corollary 2 in FL.

We suppose that the signal y is generated by stochastic process S_0 if the agent works and process S_1 if the agent shirks. This state of the appropriate process is observed at the terminal time τ , and we shall be interested in the case where τ is small. The simplest and quite standard way to do this is to assume that S_d are diffusions with common volatility σ^2 and drift $d = 0, 1$ respectively, so that the signal is distributed as $N(d\tau, \sigma^2\tau)$.

Consider first the incentive constraint

$$\rho(\tau) = \frac{q - p}{\tau} = \frac{\Phi(\bar{y} / \sigma\sqrt{\tau}) - \Phi((\bar{y} - \tau) / \sigma\sqrt{\tau})}{\tau}$$

It is easy to ensure that ρ remains bounded away from 0 as $\tau \rightarrow 0$; for example when the normalized cutoff $\bar{z} \equiv \bar{y} / \sigma \sqrt{\tau}$ is a constant independent of τ , $\lim_{\tau \rightarrow 0} \rho(\tau) = \infty$. However, with \bar{z} fixed, $p = \Phi(\bar{z})$ is a fixed positive constant, so $\mu(\tau) = p / \tau \rightarrow \infty$ and in the limit the participation constraint would be violated. Hence we must allow $\bar{z} \rightarrow \infty$ as $\tau \rightarrow 0$ to have p / τ bounded above. Thus the question becomes whether it is possible to keep p / τ bounded above at the same time allowing \bar{z} to grow sufficiently slowly that $\rho(\tau)$ remains bounded away from zero. The answer depends on the behavior of the normal distribution Φ in the upper tail where \bar{z} is large, and using bounds for the normal distribution Fudenberg and Levine [2007] show that in fact it is impossible to do so. Hence the agent cannot be induced to work when the time period is very short.

The problem with this analysis in an economic setting is that economic signals are unlikely to exactly follow a diffusion process, and in many cases will not have a continuous density when examined at a sufficiently fine scale. For example, the observed signal might be aggregate sales, which is the sum of a number of discrete random variables representing individual sales opportunities. As such we might expect from the central limit theorem that to a good approximation it follows a diffusion, and thus that the probabilities of correctly detecting a deviation and of falsely suspecting one could both be computed using the normal distribution. However, as we saw, in order to reach conclusions about incentives it is necessary to know what the probability of the signal is for very unlikely values of the normalized cutoff \bar{z} , and the central limit theorem does not help with this. Instead, to analyze a sequence of games where the signals are a sum that is approaching a diffusion limit, an extension of the central limit theorem to tail probabilities is needed.

To illustrate the usefulness of the tail probability bound, consider signals that are generated as the sums of the binomial random variables $Y_j(\Delta)$ with outcomes $\pm \sigma \Delta^{1/2}$ where the probability of the positive signal under action 0 is $\alpha_0 \equiv .5$, and the probability under action 1 is $\alpha_1 = .5 + .5\Delta^{1/2} / \sigma$. Then the expected values of the two binomials are 0 and Δ , and their standard deviations are both equal to σ . Define $k = \tau / \Delta$ to be the number of intervals of length Δ that occur during a period of length τ , where we assume that τ is an integral multiple of Δ . We take the signal to be the sum of these binomials during the period

$$y = \sum_{j=1}^k Y_j(\Delta).$$

Hence y has variance $\sigma^2\tau$ and mean either 0 or τ depending on whether the action taken is work or shirk.

To apply the central limit theorem, we should assume that the number of observations per period k grows large even as $\tau \rightarrow 0$. The key issue is that for k large while y is approximately normal it is not exactly normal, so the normal bounds used in Fudenberg and Levine [2007] do not apply directly. Moreover, Fudenberg and Levine [2007] use bounds in the upper tail of the normal, the convergence of which are not guaranteed by the central limit theorem. Hence we will need a version of the central limit theorem that applies to the tail probabilities. This in turn requires that as the period length $\tau \rightarrow 0$ the number of observations per period k grow “sufficiently fast.” In fact a sufficient condition will be $\lim_{\tau \rightarrow 0} \tau \exp(k^{2/\tau}) \rightarrow \infty$.

The best theorem we know of is the “large deviations” result of Feller [1971, pp. 548-553], which gives conditions under which the c.d.f. of normalized sums F^n satisfy

$$\frac{1 - F^n(x_n)}{1 - \Phi(x_n)} \rightarrow 1$$

as the cutoff $x_n \rightarrow \infty$. Feller’s theorem is proven and applies only in the context of the standard central limit theorem – that is, the sum of i.i.d. random variables. In our setting we are dealing with a triangular array, so we must extend Feller’s result to that case. The main part of the paper proves the relevant theorem (the “main theorem”), which gives four conditions that enable us to reach the same conclusion for triangular arrays.

The first two conditions are technical conditions on the cumulant generating function that are easily shown to be satisfied in the binomial case; see FL Lemma A.2. Thus it remains to verify the third and fourth conditions, which are that $n^{-1/6}x_n \rightarrow 0$ and $x_n \rightarrow \infty$.

In our case as we vary τ and k and implicitly Δ we will generally wish to alter the cutoff \bar{y} and the normalized version $\bar{z} \equiv \bar{y} / \sigma\sqrt{\tau}$. Suppose first that the cutoff is asymptotically very large in the sense that $\liminf_{k \rightarrow \infty} \bar{z}k^{-1/6} > 0$. Then it is shown in FL Lemma A.5 that the cutoff is sufficiently far out in the tail that there is inadequate punishment: that is $q/\tau \rightarrow 0$ (and consequently since $q \geq p$ also $p/\tau \rightarrow 0$). Hence we may assume the third condition of the main theorem. The fourth condition of the main

theorem requires $\bar{z} \rightarrow \infty$; if not, then the punishment probability does not go to zero, and as noted above this results in a trivial equilibrium. Hence we may apply the main theorem and since

$$\frac{1 - F^n(\bar{z})}{1 - \Phi(\bar{z})} \rightarrow 1$$

the normal bounds used in Fudenberg and Levine [2007] can be applied to conclude that all limit equilibria are trivial.

3. The Setup

As indicated, we extend an argument concerning i.i.d. random variables from Feller [1971, pp. 548-553] to the case of triangular arrays. We adopt Feller's notation to the maximum extent feasible. We suppose that we are given for each n a sequence Z_i^n $i = 1, \dots, n$ of i.i.d. random variables with zero mean, variance σ_n^2 and cumulative distribution function F_n . We define

$$z_n = \sum_{i=1}^n Z_i^n.$$

This has distribution F_n , while the normalized sum $z_n / \sigma_n \sqrt{n}$ has distribution F^n .

Let Φ, ϕ respectively denote the c.d.f. and density of the standard normal distribution. Recall that the *cumulant generating function*⁴ is defined as the logarithm of the moment generating function

$$\psi_n(\zeta) \equiv \log \int_{-\infty}^{\infty} e^{\zeta x} F_n(dx).$$

By the usual properties of the moment generating function, z_n has cumulant generating function $n\psi_n(\zeta)$. The derivatives of the cumulant generating function at zero are the corresponding central moments: $\psi_n'(0) = EZ_i^n$, $\psi_n''(0) = \text{var}(Z_i^n)$ and so forth. Our goal is to prove the following result:

Main Theorem: *Suppose*

1. *For some $\bar{s} > 0$ and all $0 \leq \zeta \leq \bar{s}$ there is a continuous function $\psi^2(\zeta) > 0$ and constant $B > 0$ such that*

⁴ Also called the bilateral Laplace transform.

$$\lim_{n \rightarrow \infty} \sup_{0 \leq \zeta \leq \bar{s}} |\psi_n''(\zeta) - \psi^2(\zeta)| \rightarrow 0$$

and $\qquad \qquad \qquad$ for $\qquad \qquad \qquad$ all $\qquad \qquad \qquad n$

$$\max \left\{ \sup_{0 \leq \zeta \leq \bar{s}} |\psi_n'''(\zeta)|, \sup_{0 \leq \zeta \leq \bar{s}} |\psi_n''''(\zeta)\zeta|, \sup_{0 \leq \zeta \leq \bar{s}} |\psi_n''''(\zeta)\zeta^2| \right\} < B$$

$$2. \sigma_n \rightarrow \sigma, \quad M_{3n} \equiv E|Z_i^n|^3 \rightarrow M_3 < \infty$$

$$3. n^{-1/6}x_n \rightarrow 0$$

$$4. x_n \rightarrow \infty$$

Then

$$\frac{1 - F^n(x_n)}{1 - \Phi(x_n)} \rightarrow 1$$

4. Basic Facts

The following version of the central limit theorem is taken from Feller.

Berry-Esseen Theorem.⁵ for all x

$$|F^n(x) - \Phi(x)| \leq \frac{9E|Z_i^n|^3}{\sqrt{n}\sigma_n^3}.$$

We also use some basic results about the standard normal distribution.

Lemma 1: $\lim_{x \rightarrow \infty} \frac{x^{-1}\phi(x)}{1 - \Phi(x)} = 1$

proof: This follows from l'Hopital's rule.

$$\lim_{x \rightarrow \infty} \frac{x^{-1}\phi(x)}{1 - \Phi(x)} = \lim_{x \rightarrow \infty} \frac{-(1 + x^{-2})\phi(x)}{-\phi(x)} = 1.$$

☑

Lemma 2: If Assumptions 1 and 2 of the main theorem hold then $\psi^2(0) = \sigma^2$.

⁵ Feller uses the constant 3 instead of 9; Wolfram gives 33/4 which is slightly smaller than 9. The exact value of the constant is not important in what follows.

Proof: Because it is the cumulant generating function for $Z_i^n, \psi_n''(0) = \sigma_n^2$. By Assumption 2 $\sigma_n^2 \rightarrow \sigma^2$. By Assumption 1 if $\zeta \rightarrow 0$ then $\psi_n''(\zeta) \rightarrow \psi''(0)$. But by a diagonalization argument we can then choose $\zeta \rightarrow 0$ sufficiently fast that $\psi_n''(\zeta) \rightarrow \sigma^2$. \square

5. The “Associated” Distribution

Feller’s proof replaces the normalized sum z_n and its cdf F^{n*} with a different random variable. This “associated” random variable has probability measure given by the cdf

$$V_s^{n*}(x) \equiv \int_{-\infty}^x e^{-n\psi_n(s)} e^{sy} dF^{n*}(y)$$

where s is a positive constant. The next result shows that this function integrates to 1 and so is indeed a cdf.

Lemma 3:

$$\int_{-\infty}^{\infty} e^{-n\psi_n(s)} e^{sy} dF^{n*}(y) dy = 1$$

Proof:

$$\int_{-\infty}^{\infty} e^{-n\psi_n(s)} e^{sx} dF^{n*}(x) = e^{-n\psi_n(s)} e^{\log \int_{-\infty}^{\infty} e^{sx} dF^{n*}(x)}$$

and the result follows from the fact that z_n has cumulant generating function $n\psi_n$. \square

Notice that V_s^{n*} has a thicker right tail than F^{n*} . The idea is that by applying the Berry-Esseen theorem to V_s^{n*} , we can pull this back to the thinner tailed F^{n*} to get a bound that will apply even for large values of x . First we develop some basic properties of V_s^{n*}

Lemma 4: V_s^{n*} has mean $n\psi_n'(s)$ and variance $n\psi_n''(s)$.

Proof: Follows by computing the cumulant generating function for V_s^{n*}

$$\begin{aligned} \psi_{V_s^{n*}}(\zeta) &= \log \int_{-\infty}^{\infty} e^{\zeta x} dV_s^{n*}(x) = -n\psi_n(s) + \log \int_{-\infty}^{\infty} e^{(s+\zeta)x} dF^{n*}(x) \\ &= -n\psi_n(s) + n\psi_n(s + \zeta) \end{aligned}$$

\square

Lemma 5: V_s^{n*} is the cumulative distribution function of the sum of n i.i.d. random variables with distribution

$$V_{ns}(x) \equiv \int_{-\infty}^x e^{-\psi_n(s)} e^{sy} dF_n(y)$$

Proof: This follows from the basic properties of the exponential function: multiplying a density by an exponential of the integrand commutes with the taking of convolutions.⁶

☑

6. Sketch of the Proof

We want to give a sufficient condition for

$$\left| \frac{1 - F^n(x_n)}{1 - \Phi(x_n)} - 1 \right| \rightarrow 0 \text{ as } x_n \rightarrow \infty.$$

The idea is to introduce an intermediate quantity A_n and give a sufficient condition that

$$\left| \frac{1 - F^n(x_n)}{A_n} - 1 \right| \rightarrow 0$$

and

$$\left| \frac{A_n}{1 - \Phi(x_n)} - 1 \right| \rightarrow 0,$$

the two together then giving the desired result. The first step will follow by applying the Berry-Esseen theorem to the thick tailed V^{n*} . The second step shows that when we thicken the tail by multiplying by a carefully chosen exponential we do not shift V_s^{n*} too much to the right. To carry out this second step we need the key condition $n^{-1/6}x_n \rightarrow 0$.

7. Proof of the Main Theorem

7.1. First step

Invert the relationship $dV_s^{n*}(x) = e^{-n\psi_n(s)} e^{sx} dF^{n*}(x)$ to find

⁶ Note the basic one-tailed nature of the argument: we can thicken the tail while preserving convolutions only if we multiply by an exponential. While this thickens one tail, it also thins the other tail.

$dF^{n*}(x) = e^{n\psi_n(s)}e^{-sx}dV_s^{n*}(x)$, and in particular

$$1 - F^n(x_n) = 1 - F^{n*}(x_n\sigma_n\sqrt{n}) = e^{n\psi_n(s)} \int_{x_n\sigma_n\sqrt{n}}^{\infty} e^{-sy} dV_s^{n*}(y).$$

7.2. Second step

Choose s_n to depend on n (and thus indirectly on σ_n, x_n) so that $x_n\sigma_n\sqrt{n} = n\psi_n'(s_n)$, or equivalently

$$\psi_n'(s_n) = x_n\sigma_n / \sqrt{n}.$$

Since $\lim_{n \rightarrow \infty} \sup_{0 \leq \zeta \leq \bar{s}} |\psi_n''(\zeta) - \psi^2(\zeta)| = 0$, $\psi^2(\zeta) > 0$ for all $\zeta \in [0, \bar{s}]$, $\sigma_n \rightarrow \sigma$ and $x_n / \sqrt{n} \rightarrow 0$ a solution in $[0, \bar{s}]$ exists for large enough n .

Lemma 6: *If Assumptions 1, 2, and 3 of the main theorem hold then $s_n \rightarrow 0$, $ns_n^3 \rightarrow 0$, $\psi_n''(s_n) \rightarrow \psi^2(0) = \sigma^2$, and $s_n\sqrt{n\psi_n''(s_n)} \rightarrow \infty$*

Proof: Because $\psi_n'(0) = EZ_i^n = 0$, and $\psi_n'(s_n) = x_n\sigma_n / \sqrt{n}$, by the mean value theorem we may write $\psi_n''(\zeta_n)s_n = x_n\sigma_n / \sqrt{n}$ where $\zeta_n \in [0, \bar{s}]$. Then since $x_n / \sqrt{n} \rightarrow 0$, so does

$$s_n = \frac{x_n\sigma_n}{\psi_n''(\zeta_n)\sqrt{n}}.$$

Hence $\psi_n''(s_n) \rightarrow \psi^2(0)$ by $\lim_{n \rightarrow \infty} \sup_{0 \leq \zeta \leq \bar{s}} |\psi_n''(\zeta) - \psi^2(\zeta)| \rightarrow 0$.

Now write

$$ns_n^3 = \left(\frac{n^{-1/6}x_n\sigma_n}{\psi_n''(\zeta)} \right)^3,$$

giving $ns_n^3 \rightarrow 0$. Finally

$$s_n\sqrt{n\psi_n''(s_n)} = x_n\sigma_n \frac{\sqrt{\psi_n''(s_n)}}{\psi_n''(\zeta_n)} \rightarrow \infty.$$

□

7.3. Third step

Define the quantity A_n by replacing $V_{s_n}^{n*}$ in the expression from step 1

$$e^{n\psi_n'(s_n)} \int_{x_n\sigma_n\sqrt{n}}^{\infty} e^{-s_n y} dV_{s_n}^{n*}(y)$$

by a normal with mean $n\psi_n'(s_n)$ and variance $n\psi_n''(s_n)$

$$A_n \equiv e^{n\psi_n'(s_n)} \int_{x_n\sigma_n\sqrt{n}}^{\infty} e^{-s_n y} \frac{1}{\sqrt{2\pi}\sqrt{n\psi_n''(s_n)}} e^{-(1/2)(y-n\psi_n'(s_n))^2 / n\psi_n''(s_n)} dy$$

We rewrite A_n in a more convenient form. Use the substitution

$$y = n\psi_n'(s_n) + t\sqrt{n\psi_n''(s_n)}$$

and the fact that the lower limit of integration $x_n\sigma_n\sqrt{n} = n\psi_n'(s_n)$ to find

$$A_n = e^{n[\psi_n(s_n) - \psi_n'(s_n)s_n]} \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-ts_n\sqrt{n\psi_n''(s_n)} - (1/2)t^2} dt$$

Complete the square in the numerator to get

$$A_n = e^{n[\psi_n(s_n) - \psi_n'(s_n)s_n + (1/2)\psi_n''(s_n)s_n^2]} (1 - \Phi(s_n\sqrt{n\psi_n''(s_n)}))$$

7.4. Fourth Step

Use Lemmas 4 and 5 to apply the Berry-Esseen theorem to $V_{s_n}^{n*}$ and find for all y

$$\left| V_{s_n}^{n*}(y) - \Phi\left(\frac{y - n\psi_n'(s_n)}{\sqrt{n\psi_n''(s_n)}}\right) \right| < \frac{9M_{n3}}{\sqrt{n}[\psi_n''(s_n)]^{3/2}}$$

where M_{n3} is the third absolute central moment of V_{ns_n} .⁷

7.5. Fifth Step

How close is A_n to our target $1 - F^n(x_n)$?

⁷ The parallel claim in Feller's proof is the related but different inequality $\left| V^{n*}(y) - \Phi\left(\frac{y - n\psi_n'(s)}{\sqrt{n\psi_n''(s)}}\right) \right| < \frac{9M_{n3}}{\sqrt{n}\sigma_n^3}$. This claim seems to be an incorrect application of the Berry-Esseen theorem which requires the variance of V^{n*} rather than σ_n^2 in the denominator

$$\begin{aligned}
& |1 - F^n(x_n) - A_n| = \\
& \left| e^{n\psi_n(s_n)} \int_{x_n\sigma_n\sqrt{n}}^{\infty} e^{-s_n y} dV_{s_n}^{n*}(y) - e^{n\psi_n(s_n)} \int_{x_n\sigma_n\sqrt{n}}^{\infty} e^{-s_n y} \frac{1}{\sqrt{2\pi}\sqrt{n\psi_n''(s_n)}} e^{-(1/2)(y-n\psi_n'(s_n))^2/n\psi_n''(s_n)} dy \right| \\
& = \left| e^{n\psi_n(s_n)} \int_{x_n\sigma_n\sqrt{n}}^{\infty} e^{-s_n y} \left[dV_{s_n}^{n*}(y) - \frac{1}{\sqrt{2\pi}\sqrt{n\psi_n''(s_n)}} e^{-(1/2)(y-n\psi_n'(s_n))^2/n\psi_n''(s_n)} dy \right] \right| \\
& = \left| e^{n\psi_n(s_n)} \int_{x_n\sigma_n\sqrt{n}}^{\infty} e^{-s_n y} \left[dV_{s_n}^{n*}(y) - \frac{\phi}{\sqrt{n\psi_n''(s_n)}} \left(\frac{y - n\psi_n'(s_n)}{\sqrt{n\psi_n''(s_n)}} \right) dy \right] \right|
\end{aligned}$$

Integrate by parts to find

$$\begin{aligned}
& |1 - F^n(x_n) - A_n| \leq \\
& e^{n\psi_n(s_n)} e^{-s_n x_n \sigma_n \sqrt{n}} \left| \Phi \left(\frac{x_n \sigma_n \sqrt{n} - n\psi_n'(s_n)}{\sqrt{n\psi_n''(s_n)}} \right) - V_{s_n}^{n*}(x_n \sigma_n \sqrt{n}) \right| \\
& + e^{n\psi_n(s_n)} \int_{x_n\sigma_n\sqrt{n}}^{\infty} s_n e^{-s_n y} \left| V_{s_n}^{n*}(y) - \Phi \left(\frac{y - n\psi_n'(s_n)}{\sqrt{n\psi_n''(s_n)}} \right) \right| dy
\end{aligned}$$

Now plug the bound from Step 3.

$$\begin{aligned}
& |1 - F^n(x_n) - A_n| \\
& \leq e^{n\psi_n(s_n)} \frac{9M_{n3}}{\sqrt{n}[\psi_n''(s_n)]^{3/2}} \left[e^{-s_n x_n \sigma_n \sqrt{n}} + \int_{x_n\sigma_n\sqrt{n}}^{\infty} s_n e^{-s_n y} dy \right] \\
& = \frac{18M_{n3}}{\sqrt{n}[\psi_n''(s_n)]^{3/2}} e^{n\psi_n(s_n) - s_n x_n \sigma_n \sqrt{n}} \\
& = \frac{18M_{n3}}{\sqrt{n}[\psi_n''(s_n)]^{3/2}} e^{n\psi_n(s_n) - n\psi_n'(s_n)s_n}
\end{aligned}$$

where the last step follows from $x_n\sigma_n\sqrt{n} = n\psi_n'(s_n)$. We can rewrite this as

$$\left| \frac{1 - F^n(x_n)}{A_n} - 1 \right| \leq \frac{18M_{n3}}{A_n \sqrt{n}[\psi_n''(s_n)]^{3/2}} e^{n\psi_n(s_n) - n\psi_n'(s_n)s_n}$$

Now from Step 2

$$A_n = e^{n[\psi_n(s_n) - \psi_n'(s_n)s_n + (1/2)\psi_n''(s_n)s_n^2]} (1 - \Phi(s_n \sqrt{n\psi_n''(s_n)})),$$

Plugging in on the right and rearranging terms yields

$$\left| \frac{1 - F^n(x_n)}{A_n} - 1 \right| \leq \frac{18M_{n3}}{e^{n[(1/2)\psi_n''(s_n)s_n^2]} \left(1 - \Phi(s_n \sqrt{n\psi_n''(s_n)})\right) \sqrt{n} [\psi_n''(s_n)]^{3/2}}$$

Then using the definition of ϕ we have

$$\left| \frac{1 - F^n(x_n)}{A_n} - 1 \right| \leq \frac{18M_{n3} \sqrt{2\pi} s_n (\psi_n''(s_n))^{-1} \phi(s_n \sqrt{n\psi_n''(s_n)})}{\psi_n''(s_n) (1 - \Phi(s_n \sqrt{n\psi_n''(s_n)}))}$$

From proof of lemma 6 $s_n \rightarrow 0$ so the first factor converges to 0. From Lemma 6 $s_n \sqrt{n\psi_n''(s_n)} \rightarrow \infty$ so use Lemma 1 to conclude that the second factor converges to 1 and so the entire right-hand side converges to 0.

Note that for this result we do not need $n^{-1/6}x_n \rightarrow 0$, $n^{-1/2}x_n \rightarrow 0$ would be sufficient.

7.6. Sixth and Final Step

We must now show

$$\frac{A_n}{1 - \Phi(x_n)} = e^{n[\psi_n(s_n) - \psi_n'(s_n)s_n + (1/2)\psi_n''(s_n)s_n^2]} \frac{(1 - \Phi(s_n \sqrt{n\psi_n''(s_n)}))}{1 - \Phi(x_n)} \rightarrow 1$$

We will do this by showing that both

$$e^{n[\psi_n(s_n) - \psi_n'(s_n)s_n + (1/2)\psi_n''(s_n)s_n^2]} \rightarrow 1$$

and

$$\frac{(1 - \Phi(s_n \sqrt{n\psi_n''(s_n)}))}{1 - \Phi(x_n)} \rightarrow 1.$$

7.6.1. Final Step First Half

$$e^{n[\psi_n(s_n) - \psi_n'(s_n)s_n + (1/2)\psi_n''(s_n)s_n^2]} \rightarrow 1$$

or equivalently

$$g_n(s_n) = n[\psi_n(s_n) - \psi_n'(s_n)s_n + (1/2)\psi_n''(s_n)s_n^2] \rightarrow 0.$$

Observe that $g_n(0) = 0, g_n'(0) = 0, g_n''(0) = 0$. Hence by the mean value theorem

$$g_n(s_n) = (1/6)g_n'''(\zeta)s_n^3.$$

By the uniform boundedness assumptions on the third through fifth derivatives of ψ_n , $g_n'''(\zeta)/n$ is uniformly bounded, so by Lemma 6 $g_n(s_n) \rightarrow 0$ provided $n^{-1/6}x_n \rightarrow 0$.

7.6.2. Final Step Last Half

$$\frac{(1 - \Phi(s_n \sqrt{n\psi_n''(s_n)}))}{1 - \Phi(x_n)} \rightarrow 1$$

From Lemma 6 $s_n \sqrt{n\psi_n''(s_n)} \rightarrow \infty$ and since $x_n \rightarrow \infty$, Lemma 1 implies that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{(1 - \Phi(s_n \sqrt{n\psi_n''(s_n)}))}{1 - \Phi(x_n)} \\ &= \lim_{n \rightarrow \infty} \frac{\frac{(s_n \sqrt{n\psi_n''(s_n)})^{-1} \phi(s_n \sqrt{n\psi_n''(s_n)})}{1 - \Phi(s_n \sqrt{n\psi_n''(s_n)})} (1 - \Phi(s_n \sqrt{n\psi_n''(s_n)}))}{\frac{x_n^{-1} \phi(x_n)}{1 - \Phi(x_n)} (1 - \Phi(x_n))} \\ &= \lim_{n \rightarrow \infty} \frac{x_n \phi(s_n \sqrt{n\psi_n''(s_n)})}{s_n \sqrt{n\psi_n''(s_n)} \phi(x_n)} \\ &= e^{-(1/2)[s_n^2 n \psi_n''(s_n) - x_n^2]} \frac{x_n}{s_n \sqrt{n\psi_n''(s_n)}} \\ &= e^{-(1/2)[s_n^2 n \psi_n''(s_n) - n[\psi_n'(s_n)]^2 / \sigma_n^2]} \frac{\psi_n'(s_n)}{\sigma_n s_n \sqrt{\psi_n''(s_n)}} \end{aligned}$$

Consider first by the mean value theorem for some $0 \leq \zeta_n \leq s_n$

$$\frac{\psi_n'(s_n)}{\sigma_n s_n \sqrt{\psi_n''(s_n)}} = \frac{\psi_n''(\zeta_n) s_n}{\sigma_n s_n \sqrt{\psi_n''(s_n)}} \rightarrow \frac{\psi''(0)}{\sigma \sqrt{\psi''(0)}} = 1$$

where we apply Lemma 6 to find the limit. So we are left with showing

$$h_n(s_n) = n[s_n^2 \psi_n''(s_n) - [\psi_n'(s_n)]^2 / \sigma_n^2] \rightarrow 0.$$

Here $h_n(0) = 0, h_n'(0) = 0, h_n''(0) = 0$ so for some $0 \leq \zeta_n \leq s_n$

$$h_n(s_n) = (1/6)h_n'''(\zeta_n)s_n^3.$$

Again by the uniform boundedness assumptions on the third through fifth derivatives of ψ_n , $h_n'''(\zeta)/n$ is uniformly bounded, so by Lemma 6 this is true provided $n^{-1/6}x_n \rightarrow 0$.



References

- Feller, W. [1971] *An Introduction to Probability Theory and Its Applications*, Volume II, John Wiley & Sons, New York.
- Fudenberg D. and D. K. Levine [2007] “Continuous Time Limits of Repeated Games with Imperfect Public Monitoring,” *Review of Economic Dynamics*, 10: 173-192
- Fudenberg D. and D. K. Levine [2009] “Repeated Games with Frequent Signals,” *Quarterly Journal of Economics* **124**: 233-265.
- Mirlees, J. A. [1974] “Notes on Welfare Economics, Information and Uncertainty,” in M. Balch, D. McFadden and S.-Y. Wu, editors, *Essays on Economic Behavior under Uncertainty*, North Holland, Amsterdam.
- Sannikov, Y. [2007] “Games with Imperfectly Observed Actions in Continuous Time,” *Econometrica* 75, 1285-1329.
- Sannikov, Y., and A. Skrzypacz [2007] “Impossibility of Collusion under Imperfect Monitoring with Flexible Production,” *American Economic Review* 97, 1794-1823.
- Sadzik, T. and E. Stacchetti [2012] “Agency Models with Frequent Actions: A Quadratic Approximation Method,” mimeo.